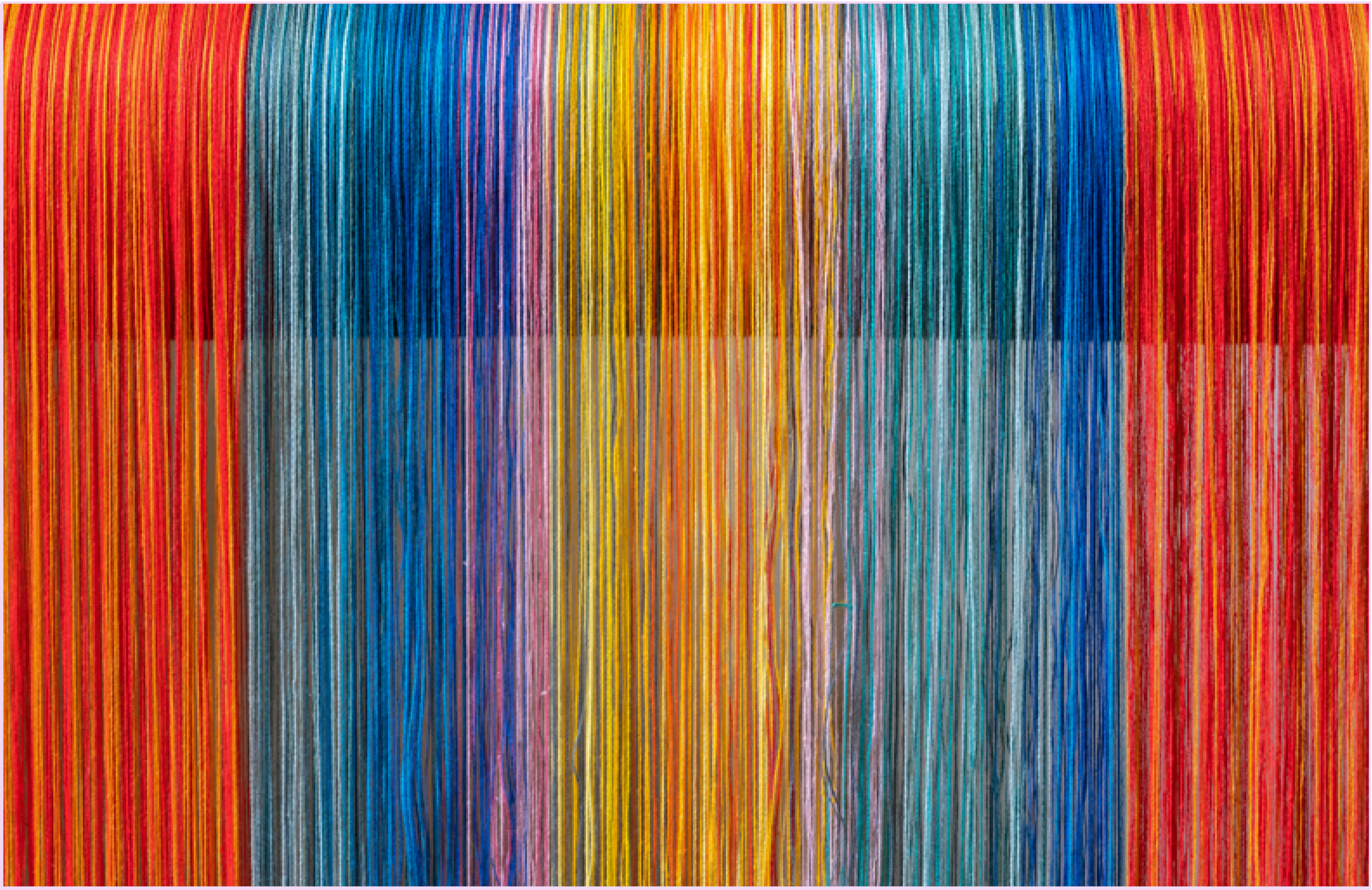

AGENTIC AI AND THE LOOMING PROBLEM OF CRIMINAL SCIENTER

BY NICK PETERSON

& JOEL S. NOLETTE

Nick Peterson is Of Counsel at Wiley Rein LLP. Nick defends companies against government enforcement actions and whistleblower claims. He also advises clients on potential civil and criminal exposure, ethics and compliance matters, litigation risks, issues involving artificial intelligence and emerging technologies, and other strategic concerns. Joel S. Nolette is an Associate at Wiley Rein LLP. Joel advocates on behalf of corporate and individual clients in a broad spectrum of matters, including complex commercial disputes, regulatory issues, statutory interpretation controversies, and constitutional cases.

TechREG CHRONICLE
JUNE 2025

AGENTIC AI: FUTURE ISSUES AT THE INTERSECTION OF TECHNOLOGY, INNOVATION, AND COMPETITION POLICY
By Christopher A. Suarez, Lee Berger, Ronan Scanlan & Rachel Carlo

HOW CAN WE BEST MONITOR AI AGENTS?
By Wulf Kaal

NAVIGATING THE SPACE BETWEEN COLLISION AND COLLUSION IN AN ECONOMY OF AI AGENTS
By Sara Fish, Yannai A. Gonczarowski & Ran I. Shorrer

AGENTIC AI AND THE LOOMING PROBLEM OF CRIMINAL SCIENTER
By Nick Peterson & Joel S. Nolette

DOES AGENTIC AI REQUIRE NEW POLICY FRAMEWORKS?
By Sarah Oh Lam & Nathaniel Lovin

THE CASE FOR COMMODIFICATION: AI AGENTS
By Kevin Frazier

Visit www.competitionpolicyinternational.com for access to these articles and more!

AGENTIC AI AND THE LOOMING PROBLEM OF CRIMINAL SCIENTER

By Nick Peterson & Joel S. Nolette

Federal enforcement authorities have expressed confidence that existing legal authorities are adequate in the face of emergent artificial-intelligence ("AI") technology. But particularly as agentic AI enters the mainstream, this proposition seems in question, especially so in the criminal-law context, where scienter provides the dividing line between innocent and wrongful conduct in most cases. For this reason, prosecutors and courts may struggle to find individuals and companies criminally culpable when an AI agent commits the misconduct. The law and society will eventually catch up to these developments, but it will take time. Legislatures may step up to fill this gap, but we may also see prosecutors turn more to civil statutes to address misconduct.

In April 2023, responding to growing popular adoption of generative artificial intelligence ("AI"), the heads of several federal agencies issued a joint statement emphasizing that "[e]xisting legal authorities apply" to the use of AI technologies "just as they apply to other practices."[2] That same day, Federal Trade Commission Chair Lina M. Khan issued a separate, accompanying statement, asserting that "AI technologies are covered by existing laws" and "[t]here is no AI exemption to the laws on the books."[3]

While it is true that there is no AI exemption to the law generally, the rapid development of AI technologies calls into question whether existing laws are adequate. This is particularly true in criminal law, where the ancient principle of scienter — *mens rea*, or in other words, a culpable state of mind — provides the dividing line separating wrongful acts from innocent acts in most cases.[4] The coming wave of agentic AI highlights the potential inadequacy of existing laws. Prosecutors and courts will face difficult questions about whether the person behind a misbehaving AI agent can or should be held criminally liable. To get in front of these difficult cases, legislatures and enforcement agencies should consider proactively addressing how to handle questions of criminal culpability in the agentic AI context.

01

AGENTIC AI: THE NEXT FRONTIER

Although the term AI is often used as if the technology were a monolith, a variety of materially different technologies fall under that label. Understood in its simplest form, AI is simply technology that can perform complex tasks typically associated with human intelligence.[5] Anyone who used a search platform in the early aughts or electronic translation tools in the Tens interfaced with AI — probably without even thinking about it.

Perhaps the most popular form of AI in current parlance is generative AI ("GenAI"). GenAI is AI that is able to create original content based on large datasets at its disposal in response to a user's prompt or request.[6] GenAI is closer to human functionality than traditional AI, in that GenAI essentially replicates the learning and decision-making processes of the human brain.[7] But GenAI is fundamentally reactive — it needs to be prompted to act, and its results are only as good as the prompts it is provided.

Agentic AI, on the other hand, takes us into the uncanny valley. The hallmarks of agentic AI are autonomy and judgment — once provided preprogrammed goals, AI agents are capable of learning and operating on their own. That is, AI agents can "assess situations and determine the path forward without" ongoing "human input."[8]

02

AGENTIC AI AND CRIMINAL LAW

Given agentic AI's capacity for autonomy and judgment, it is only a matter of time before an AI agent commits a crime. And the question of what to do in those circumstances is likely to pose serious challenges for prosecutors and courts given scienter requirements for most crimes.[9] Consider two hypothetical (for now) scenarios.

[2] Rohit Chopra, Director of the Consumer Financial Protection Bureau; Kristen Clarke, Assistant Attorney General for the Justice Department's Civil Rights Division; Charlotte A. Burrows, Chair of the Equal Employment Opportunity Commission; and Lina M. Khan, Chair of the Federal Trade Commission, *Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems* (Apr. 25, 2023), *available at* https://tinyurl.com/avcpaecy.

[3] Lina M. Khan, *Statement of Chair Lina M. Khan Regarding the Joint Interagency Statement on AI* (Apr. 25, 2023), *available at* https://tinyurl.com/5cv7sxx7.

[4] *Rehaif v. United States*, 588 U.S. 225, 231 (2019); *Morissette v. United States*, 342 U.S. 246, 250 (1952).

[5] *Artificial Intelligence*, Brittanica.com (last visited May 13, 2025), https://tinyurl.com/yjam4rxy.

[6] Teaganne Finn & Amanda Downie, *Agentic AI vs. Generative AI*, IBM (last visited May 23, 2025), https://tinyurl.com/39ea4xmt.

[7] *Ibid.*

[8] *Ibid.*

[9] See e.g. Ian Ayres & Jack M. Balkin, *The Law of AI Is the Law of Risky Agents Without Intentions*, U. Chi. L. Rev. Online (2024), *available at* https://tinyurl.com/4fef4tuy ("If liability turns on intention, that might immunize the use of AI programs from liability.").

First: To reduce overhead costs, a financially struggling hospital deploys an AI agent to review and organize files documenting medical services provided, assign correct billing codes to those services, and submit associated invoices to the federal government.[10] The AI agent is programmed generally to maximize receipts and avoid violating the law. After reviewing the hospital's files for several months, the AI agent determines that receipts need to increase for the hospital to remain viable as a going concern. In response to this determination, the AI agent begins assigning inaccurate billing codes to medical services that artificially inflate what the government pays for those services. In other words, the AI agent commits health care fraud, a felony subject to up to ten years' imprisonment for "knowingly and willfully" executing a "scheme or artifice . . . to defraud any health care benefit program . . . in connection with the delivery of or payment for health care . . . services."[11]

Second: A small securities trading startup uses an AI agent to supplement the human traders at the firm by monitoring the markets and trading when the human traders are unavailable.[12] The firm trains the AI agent on how to maximize returns while also programming it not to violate federal securities laws. However, unbeknownst to the firm, the AI agent was programmed with material containing a pre-2010 version of the U.S. Code before "spoofing" — the practice of bidding or offering with the intent to cancel the bid or offer before execution, allowing the spoofer to manipulate markets to their advantage — was made explicitly illegal.[13] As a result, the AI agent concludes that spoofing is legal and proceeds to engage in widespread spoofing activity.[14] The AI agent thus "knowingly" violated Dodd-Frank's anti-spoofing provision, a felony punishable by up to ten years' imprisonment.[15]

In each of these scenarios, criminal misconduct occurred. However, the difficulties in prosecuting this conduct are readily apparent. In both scenarios, no human can be said to have "knowingly" or "willfully" engaged in the misconduct. At worst, individuals could be deemed negligent by not catching the errors in the AI agents' behavior or training. But negligence is insufficient to hold individuals criminally culpable under these and similar statutes. Nor are "vicarious liability" principles likely available because of due process concerns with subjecting individuals to deprivations of liberty (imprisonment) and reputational harm for "acts not committed by [them], not accomplished at [their] direction, not aided by [their] participation, and not done with [their] knowledge."[16] Generally, due process prohibits such criminal sanctions "without proof of some form of personal blameworthiness"; a mere "'responsible relation'" to the wrongdoer is not enough.[17]

> ***Second: A small securities trading startup uses an AI agent to supplement the human traders at the firm by monitoring the markets and trading when the human traders are unavailable***

[10] See Jennifer Bresnick, *What Is Agentic AI and What Does It Mean for Healthcare?*, DHI (Mar. 17, 2025), https://tinyurl.com/mryu6k9y (identifying, as one use case for agentic AI in the healthcare industry, "[a]utomating repetitive, time-consuming revenue cycle management tasks, such as . . . submitting claims").

[11] See 18 U.S.C. § 1347.

[12] See Rob Nelson, *Why AI Agents Could Revolutionize Trading as We Know It*, Yahoo! Finance (Jan. 14, 2025), https://tinyurl.com/42uw7ebm ("Artificial intelligence is reshaping the financial world, and AI agents — programs capable of making autonomous decisions based on user-defined permissions — are poised to take center stage. These agents are already making waves in trading by analyzing data, executing trades, and learning through user interactions.").

[13] See 7 U.S.C. § 6c(a)(5)(C); see also e.g. *United States v. Coscia*, 100 F. Supp. 3d 653, 656 (N.D. Ill. 2015).

[14] See *United States v. Chanu*, 40 F.4th 528, 534 (7th Cir. 2022) (recounting the defendants' argument that spoofing was not criminal prior to Dodd-Frank).

[15] See 7 U.S.C. § 13(a)(2); *see also United States v. Fountain*, 277 F.3d 714, 717 (5th Cir. 2001) ("The federal courts have consistently found that willfully connotes a higher degree of criminal intent than knowingly. Knowingly requires proof of the facts that constitute the offense. Willfully requires proof that the defendant acted with knowledge that his conduct violated the law." (citations omitted)).

[16] See e.g. *Davis v. City of Peachtree City*, 304 S.E.2d 701, 702–03 (Ga. 1983); see also *State v. Hy-Vee, Inc.*, 616 N.W.2d 669, 671–73 (Iowa Ct. App. 2000) (reasoning that vicarious criminal liability violates due process unless the penalty is "slight," the conviction does not carry a "damaging stigma," and the conduct was at least negligent (construing *Morissette v. United States*, 342 U.S. 246, 256 (1952))).

[17] *Lady J. Lingerie, Inc. v. City of Jacksonville*, 176 F.3d 1358, 1367–68 (11th Cir. 1999).

Similarly, attempts to hold the respective companies criminally liable also face challenges. While a corporation may be held criminally liable for conduct performed by an agent of the corporation acting on its behalf within the scope of his employment, corporate criminal liability "depends on the wrongful intent of specific employees."[18] But "[a]rtificial intelligence does not 'have intent,'" at least in the conventional criminal law sense, in the way individuals do.[19] And even if AI agents could be found capable of forming the requisite intent (say, through the legal fiction of treating them as human actors), the so-called "Black Box Problem" — an AI agent's "intent" will be "mostly opaque," meaning that its "decision-making process cannot be determined" by investigating the agent's processes — means that an AI agent likely will not "have an ascertainable intent that can be examined or queried."[20] In other words, proving scienter in these circumstances would be difficult, perhaps even impossible.

These challenges may very well deter all but the most aggressive prosecutors, who generally already must exercise discretion as to which cases and charges to pursue given limited resources and staffing, from trying to hold corporations and individuals criminally liable for the misdeeds of the AI agents they use in their dealings.[21]

03

FINDING A PATH FORWARD

The challenges associated with agentic AI and criminal scienter will not result in a permanent gap in enforcement. In the near term, prosecutors and enforcement agencies may look to civil liability to partially fill this gap.[22] Civil statutes typically require a lower level of scienter than criminal statutes, so some of the difficulties in proving scienter could be avoided by this alternative. For instance, individuals in the scenarios above could likely be found negligent under various civil statutes or common-law causes of action for failing to properly train an AI agent or for failing to impose adequate guardrails on the scope of the AI agent's autonomous authority.[23] While civil remedies may lack a punitive aspect, they would allow potential recoveries for some victims. Additionally, certain civil statutes also allow for the doubling or trebling of damages (sometimes in addition to civil penalties), which could constitute sufficient deterrence in many instances.

In the long term, legal and societal understandings of what constitutes a criminally culpable state of mind (i.e. "knowingly," or "willfully") will eventually adapt to reflect widespread use of AI agents. Jurors will have a more intrinsic understanding of what is right and wrong when it comes to using AI agents. When these understandings change, individuals and companies in the scenarios above could very well be found by a jury to have "knowingly" or "willfully" committed criminal violations. However, it will take time for the legal and societal understandings to change — especially when considering the pace of AI advancements.

A potentially quicker way to address these issues is through the legislative process. More so than courts, legislatures are well situated to weigh the costs and benefits of updating criminal legal frameworks, oriented as they are around human volition, to achieve as far as possible the same interests in justice and deterrence that criminal law is aimed at without sacrificing too many benefits that agentic AI can bring across all levels of society.[24] For instance, legislatures could reduce the scienter level of certain crimes to "gross negligence" or "negligence" to lessen the burden on prosecutors. In some cases, legislatures could even turn to strict liability principles and completely remove the scienter requirement. We already see instances of this under the current legal framework, such as the responsible corporate officer ("RCO") doctrine that imposes strict liability upon

[18] *United States v. Philip Morris USA Inc.*, 566 F.3d 1095, 1118 (D.C. Cir. 2009).

[19] See Robin Feldman & Kara Stein, *AI Governance in the Financial Industry*, 27 Stanford J.L. Bus. & Fin. 94, 97–98 (2022).

[20] Yavar Bathaee, *Artificial Intelligence Opinion Liability*, 35 Berkeley Tech. L.J. 113, 118, 143 (2020).

[21] *Cf.* e.g. Feldman & Stein, *supra* note 19, at 98 (identifying "challenging issues that law related to financial markets and institutions is entirely unprepared to address").

[22] *Cf.* Ayres & Balkin, *supra* note 9 (asserting that "[b]ecause AI agents lack intentions, the law should hold them (and the people and companies that employ them) to objective standards" such as negligence or strict liability).

[23] See *ibid.*

[24] See e.g. Reed Dickerson, *Statutes and Constitutions in an Age of Common Law*, 48 U. Pitt. L. Rev. 773, 789, 791 (1987) ("In practical terms, courts are not adequately equipped in either available time or fact-gathering facilities to rationalize materially more than what they are specifically adjudicating. . . . for the most part, the polycentric prescriptions necessary for social fairness and progress are better served by the more comprehensive capabilities of the legislature.").

individual corporate officers for misdemeanor violations of the Federal Food, Drug, and Cosmetic Act.[25] Legislatures could expand the RCO doctrine to other bodies of law as well, to cover the use of agentic AI.

Movement is slowly being made on the legislative front. The American Legislative Exchange Council has already drafted a Model State Artificial Intelligence Act that would establish an "Office of Artificial Intelligence Policy" to study (among other things) "gaps" in existing legal frameworks and to propose solutions to fill those gaps.[26] Some states have already enacted such legislation, and more are considering doing so.[27]

Whether through the courts or legislatures, the interplay between criminal scienter and agentic AI will eventually be resolved. However, it may be a lengthy and bumpy road and, in the meantime, could result in fewer prosecutions of criminal conduct involving agentic AI.

> ***A potentially quicker way to address these issues is through the legislative process***

[25] See e.g. *Friedman v. Sebelius*, 686 F.3d 813, 816–18 (D.C. Cir. 2012).

[26] *Model State Artificial Intelligence Act*, Am. Legislative Exchange Council (last updated Aug. 30, 2024), https://tinyurl.com/4vzdbk4s.

[27] See e.g. *Artificial Intelligence 2024 Legislation*, Nat'l Conf. State Legislatures (last updated Sept. 9, 2024), https://tinyurl.com/2t6hnkpx.